# Afrikaans in a quantitative typology of Germanic standard and non-standard varieties

**Simon Pröll** (University of Freiburg, Freiburg, Germany)
*simon.proell@germanistik.uni-freiburg.de*

**Thilo Weber** (IDS Mannheim, Mannheim, Germany)
*weber@ids-mannheim.de*

## Abstract

Besides English, Afrikaans is considered "the [Germanic] language which deviates grammatically the farthest from the others" (Harbert 2007: 17). But how exactly do we measure "grammatical deviation", and how deviant is Afrikaans really if we compare it not just to other standard languages but also to non-standard varieties? The present contribution aims to address those questions combining functional-typological and dialectometric perspectives. We first select data for 28 Germanic varieties showing vastly different speaker numbers, grades of standardisation and amounts of language contact. Based on 48 (micro)typological variables from syntax, morphology and phonology, we perform cluster analysis and multidimensional scaling and present ways of visualizing and interpreting the results. Inter alia, the analyses show a major divide between Continental West Germanic and North Germanic (as might be expected) and they also identify a number of outliers, including English and pidgin and creole languages such as Russenorsk or Rabaul Creole German. Afrikaans appears to cluster with the other West Germanic languages rather than the outliers. Within West Germanic, however, it does indeed emerge as rather deviant and, according to our metric, it is, for example, typologically closer to other high-contact varieties such as Yiddish than it is to Dutch.

Keywords: Afrikaans, dialectometry, Germanic, typology, variation, cluster analysis, multidimensional scaling, visualisation, phonology, morphology, syntax

## Opsomming

*Afrikaans in 'n kwantitatiewe tipologie van Germaanse standaard en nie-standaard variëteite*

Afgesien van Engels, word Afrikaans beskou as "the [Germanic] language which deviates grammatically the farthest from the others" (Harbert 2007: 17). Maar hoe kan ons "grammatikale afwyking" meet, en hoe afwykend is Afrikaans werklik as ons dit ook met niestandaardvariëteite vergelyk (nie net met ander standaardtale nie)? Hierdie artikel poog om daardie vrae aan te spreek met behulp van funksioneel-tipologiese en dialektometriese perspektiewe te kombineer. Ons kies eers data vir 28 Germaanse variëteite wat baie verskillende sprekergetalle, grade van standaardisering en hoeveelhede taalkontak toon. Gebaseer op 48 (mikro)tipologiese sintaksiese, morfologiese en fonologiese veranderlikes voer ons trosanalise en multidimensionele skalering uit en wys maniere om die resultate te

**How to cite this article:**
Pröll, Simon & Thilo Weber. 2023. "Afrikaans in a quantitative typology of Germanic standard and non-standard varieties". In *Proceedings of the 4th International Afrikaans Grammar Workshop,* edited by Adri Breed. Potchefstroom: North-West University. pp. 39-68. DOI:10.25388/nwu.25052702

visualiseer en te interpreteer. Die ontledings toon onder meer 'n groot skeiding tussen Kontinentale Wes-Germaans en Noord-Germaans (soos verwag kan word) en dit identifiseer ook 'n aantal uitskieters, insluitend Engels en pidgin en kreoolse tale soos Russenorsk of Rabaul Kreools Duits. Afrikaans groepeer blykbaar met die ander Wes-Germaanse tale, nie met die uitskieters nie. Maar binne Wes-Germaans kom dit inderdaad as taamlik afwykend na vore. Volgens ons metriek is Afrikaans byvoorbeeld tipologies nader aan ander hoëkontakvariëteite soos Jiddisj as wat dit aan Nederlands is.

Sleutelwoorde: Afrikaans, dialektometrie, Germaans, tipologie, variasie, trosanalise, multidimensionele skalering, visualisering, fonologie, morfologie, sintaksis

## 1 Introduction

Besides English, Afrikaans is considered “the [Germanic] language which deviates grammatically the farthest from the others” (Harbert 2007: 17). An example from the nominal system is Afrikaans’ total elimination of grammatical gender, shared by only one other Germanic standard language, viz. English (Harbert 2007: 100). An example from verbal morphology is its loss of a distinction between the class of so-called “strong” verbs (showing ablaut) and the class of so-called “weak” verbs (displaying a dental suffix) (Donaldson 1993: 223), which is possibly unique to Afrikaans (Harbert 2007: 277). But how exactly do we measure “grammatical deviation” more globally, and how deviant is Afrikaans really if we compare it not just to other standard languages but also to non-standard varieties (including other high-contact-varieties) of Germanic? For example, while English and Afrikaans are the only Germanic **standard** languages that have lost grammatical gender, the feature is, at least sporadically, also found in non-standard varieties such as e.g. Western Danish dialects (Harbert 2007: 100–101, Braunmüller 2000: 28–29) and a local variety of Low German (Dützmann 1939: 1–5). The present contribution aims to address those questions by bringing together two perspectives:

- Functional typology (cf. e.g. Dahl 2000 on future tense): A high level of abstraction is deliberately chosen for this; similarities and differences are assessed from a purely functional perspective. Typology is traditionally based on data from unrelated languages, and when it comes to European languages, it usually focuses on their respective standard varieties (see e.g. Murelli and Kortmann 2011, Fleischer 2015: 203, Seiler 2019). With respect to the methods of data collection, typological studies are often based on secondary data such as reference grammars.
- Quantitative dialectology or dialectometry: Dialectometry yields a microtypological classification of non-standard varieties of a single language (e.g. Lameli 2013 for the German speaking area); it relies on primary data and focuses on formal differences (see e.g. Kehrein 2009), often in large (aggregated) quantities of features.

To be sure, there is also an increasing body of research that combines dialectology and typology, including areal typology (e.g. Kortmann ed. 2004, Fleischer 2015, Seiler 2019). The (Electronic) World Atlas of English (Kortmann, Lunkenheimer & Ehret eds. 2020), for example, is a database on morphosyntactic variation in English designed for immediate comparison with typological databases such as WALS (World Atlas of Language Structures) and APiCs (Atlas of Pidgin and Creole Structures Online).

What still appears to remain a desideratum, however, is a combination of typology and dialectometry across individual languages and their varieties – i.e., an approach that would allow us to compare Afrikaans not just to e.g. Standard German, Standard Dutch and Standard Swedish but also to their respective non-standard varieties such as Bavarian, Limburgish and Elvdalian.[1] A combination of the two perspectives appears particularly fruitful for the study of Afrikaans, whose status among the Germanic varieties has been discussed for most of its existence, with preconceptions about its origin (see e.g. Roberge 2010 for an overview) often obstructing an objective assessment. To this end, we first select appropriate data for a sample of 28 Germanic varieties that show vastly different speaker numbers, grades of standardisation and amounts of language contact. On the basis of a total of 48 (micro)typological variables from syntax, morphology and phonology, we perform quantitative procedures stemming from dialectometry, namely cluster analysis and multidimensional scaling, and we present ways of visualizing and interpreting the results. Our study is led by two main questions (or rather, sets of questions): an abstract, methodological one (1) and a concrete, philological one (2):

1. (How) can we measure grammatical similarity/dissimilarity?
2. How (dis-)similar is Afrikaans to other varieties of Germanic if we consider not only standard varieties but also non-standard varieties, including other high contact varieties such as e.g. Yiddish?

In addition to the specific question of the typological status of Afrikaans among Germanic, in the long run, the approach envisaged in this paper may contribute to more general questions such as the relative role played by individual factors known and/or hypothesised to be responsible for (dis)similarity, such as:

- degree of genetic relatedness: Different varieties of Germanic are related to one another to different degrees. According to the standard family tree (see e.g. Harbert 2007: 8), the major division is that between North Germanic and West Germanic.
- amount and patterns of language contact: Trudgill (e.g. 2011: 64), for example, argues that there is a connection between contact/isolation and structural complexity: "It is therefore worth pursuing the hypothesis that it is in low contact communities that we are most likely to find not only preservation of complexity but also an increase in complexity, i.e. irregularity, opacity, syntagmatic redundancy and non-borrowed morphological categories"
- degree of standardisation: Chambers (e.g. 2004: 128), for example, proposes the notion of "vernacular universals", according to which "a small number of phonological and grammatical processes recur in vernaculars wherever they are spoken", while Chomsky (1995: 51) points out that standard languages can be viewed as "partially invented" and might even (in stark contrast to natural non-standardised languages) „violate the principles of language".
- 'roofing' by the same standard variety ("Dachsprache") or different standard varieties: Smits (2011), for example, shows how the closely related varieties spoken west and

[1] Given the particular philological question we aim to address in this paper (i.e. how different is Afrikaans compared to other Germanic languages? See below.), we apply this approach to languages all taken from the same family. In principle, however, it may just as well be applied to (varieties of) unrelated languages.

east of the Dutch-German border have gradually been converging towards Standard Dutch and Standard German, respectively.

## 2 Methods and data

In order to take the first quantitative steps towards an answer to the questions outlined in the introduction, we compiled a sample consisting of Afrikaans and 27 other varieties of Germanic, including both standard and non-standard varieties (section 2.1). We then attempted to classify them with respect to a number of grammatical (phonological, morphological and syntactic) features (section 2.2). As is customary in language typology, our classifications are based on secondary data such as reference grammars. To be sure, this way of data collection is not without its problems and limitations, which will therefore be critically reflected on (section 2.3). Based on the sample, we ultimately aimed to determine patterns of similarity and dissimilarity among the varieties represented using methods commonly used in dialectometry (presented in section 2.4).

### *2.1 Language sample*

The 28 varieties making up our sample are listed alphabetically in Table 1. We deliberately chose languages that differ widely in speaker size and sociolinguistic background. Some of them are standard languages with a large number of speakers, codified writing systems and official (political) status (e.g. English, German, Swedish)[2], while others are vernaculars with (sometimes drastically) fewer speakers and without similar official status such as Yiddish, traditional dialects spoken in Europe (e.g. Gutnish, Limburgish, Zurich German), language-island-varieties (e.g. Pomerano) as well as pidgin and creole varieties (e.g. Pitkern-Norf'k, Rabaul Creole German).

[2] See the volume by Deumert & Vandenbussche (2010) (eds.), a comparative overview of standardization in Germanic languages. Each chapter deals with an individual language or group of languages, and all chapter authors were asked to structure their chapters according to the four-step model of standardization developed by Haugen (1966 et seq.), thus allowing readers to assess the degree of standardization of the different languages relative to one another.

*Table 1: Germanic varieties considered in this study (alphabetical order).*

| variety | genealogical descent | variety | genealogical descent |
|---|---|---|---|
| Afrikaans | West Germanic | North Frisian | West Germanic |
| Basel German | West Germanic | Northern Low German | West Germanic |
| Central Hessian | West Germanic | Norwegian | North Germanic |
| Danish | North Germanic | Pennsylvania German | West Germanic |
| Dutch | West Germanic | Pitkern-Norf'k | West Germanic / Polynesian |
| Elvdalian | North Germanic | Pomerano | West Germanic |
| English | West Germanic | Rabaul Creole German | West Germanic / Austronesian |
| Faroese | North Germanic | Russenorsk | North Germanic / Slavic |
| German | West Germanic | Saterland Frisian | West Germanic |
| Gutnish | North Germanic | Swedish | North Germanic |
| Icelandic | North Germanic | Ulster Scots | West Germanic |
| Limburgish | West Germanic | West Frisian | West Germanic |
| Luxembourgish | West Germanic | Yiddish | West Germanic |
| Mòcheno | West Germanic | Zurich German | West Germanic |

### *2.2 Features*

We looked at a total of 48 features, all listed in Table 2: 16 morphological, 20 syntactic, 10 phonological and 2 'other' features that could be classified as lexical and/or historical. See appendix A for an illustration of each feature.

*Table 2: Structural features considered in this study.*

| morphological features | syntactic features | phonological features |
|---|---|---|
| gender marking on the definite article | pro-drop possible | lexical tone |
| case marking on the definite article | pronoun doubling for emphasis | location of stress |
| (loss of) derivational prefixes | verb second | front rounded vowels |
| number distinction in second person pronouns | basic order of verb and object | diphthongs |
| dual category in pronouns | doubly-filled COMP | phonological contrast through nasalized vowels |
| subject-verb-agreement | extended use of the present perfect | distinct vowel quantities |
| syncretism of verbal singular forms | 'be' as a perfect auxiliary | long consonants |
| syncretism of verbal plural forms | double perfect | velar nasal |
| 'ge'-prefix on past participles (in perfect constructions) | passive marking with 'become' | number of series of fricatives |
| past subjunctive forms (distinct from past indicative) beyond 'be' | mediopassive from old reflexive | type of contrast between the two series of plosives |
| loss of the synthetic preterite | inherent reflexivity marking | |
| (new) synthetic passive voice | highly grammaticalised progressive | |
| presence of an infinitive suffix | highly grammaticalised future | **'other' features** |
| (reflex of the) distinction between strong and weak adjective inflection | modal verb as a future auxiliary | directional adverbs |
| subject-agreement marking on predicative adjectives | distinct habitual construction | inherited /ɣ/ |
| morphological number marking on nouns | presence of an indefinite article | |
| | presence of a definite article | |
| | posievation of the definite article | |
| | external possessors | |
| | multiple negation | |

Concerning the selection of these features, our starting point was McWhorter's (2002, 2004) approach to the question "What happened to English?", which contains 10 features common to all or most Germanic languages but absent from English. We adopted nine of those ten

features that may be revealing when it comes to distinguishing typical and atypical Germanic languages.[3] For additional 39 features, we consulted comparative descriptions of the Germanic languages (most notably Harbert 2007), the WALS database as well as (areal-) typological studies (for standard languages), e.g. Thieroff (2000) on tense and aspect in European languages, and reference grammars of individual varieties. Our classification of Afrikaans is, for the most part, based on the reference grammar by Donaldson (1993). Overall, the selection of features was guided by the availability of data for a decent number of languages/varieties from our sample and the representation of different levels of language structure: phonology, (inflectional) morphology and syntax. Still, inevitably, a large element of randomness remains.

### *2.3 Challenges for our feature-based approach*

Additionally, there is the issue of trying to express complex data situations in terms of quantifiable features with a relatively small number of feature values. This often comes with certain problems (see e.g. Seiler 2019: 545–548, Schallert, Seiler & Weber under review). First of all, with some features, the value to be assigned to a given language/variety is controversial or theory-dependent. A prominent example is the order of verb and object in Continental West Germanic. According to WALS (Dryer 2005), for example, German, Dutch and Frisian have "no dominant order" of subject, object and verb. According to Haider (2013: 98), by contrast, the Continental West Germanic languages, including Afrikaans, have "OV sentence structure". In cases like those, our approach was to keep the potential systematic error constant, i.e. to choose one framework and apply it consistently (in this particular case we followed Haider).

In some cases, the feature value for a given variety depends on the age and/or conservativeness of the source that is consulted. For example, according to one of the major reference grammars of Northern Low German, Lindow et al. (1998: 284–285), the variety in question has multiple negation, more precisely both negative doubling (combining verbal negation with a negative indefinite) and negative spread (combining multiple negative indefinites). A more recent study (Elmentaler & Borchert 2012: 122), however, finds that the relevant negation patterns are apparently no longer used in present-day Low German. In cases like those, we followed the more recent description. Nonetheless, the example illustrates a general problem: The data for our study had to be drawn from fairly diverse sources and has an inherent diachronic spread, with a range from ca. 1950 to 2020.[4]

What is more, very often a given feature is not simply present or absent but there are levels in between. An example is the feature "loss of the synthetic preterite". While in some languages, the preterite has been lost completely (e.g. Zurich German, cf. Weber 1948: 194–195) or almost completely (with residual forms of e.g. the copula or modal verbs as in Afrikaans, cf. Donaldson 1993: 223), other languages such as Pomerano (Postma 2014: 655–659) or Central Hessian (Hasselberg 1979: 120) also show fairly extensive preterite loss but do retain preterite forms for more than just a handful of verbs. In the variety of Pomerano described by Postma (2014: 655–659) for example, the preterite has been lost for the class of weak verbs

[3] At the same time, it means that our feature selection has a slight bias towards singling English out.
[4] Russenorsk is an outlier when it comes to the diachronic dimension of our sample, in the sense that it has been extinct for about one century now (cf. Jahr 2005: 1538). All other varieties have L1-speakers at the time of writing this article.

(which traditionally formed their preterite by means of the so-called dental suffix) but is retained for the class of strong verbs (forming their preterite by means of ablaut). In this particular case as well as some comparable ones, we operated with ternary rather than binary feature values. In the case of preterite loss, we assigned the value 0 where the preterite is preserved (e.g. English, Swedish, Standard German), we assigned the value 1 where the preterite is completely or almost completely lost (Zurich German, Afrikaans), and we assigned the value of 0.5 for intermediate cases such as Pomerano or Central Hessian.

Still, there remain features where finer distinctions would be appropriate than we were able to make: For grammatical gender marking on the definite article, for example, we only made a binary distinction (viz "yes", e.g. Danish, Dutch, West Frisian, Icelandic, vs. "no", e.g. English, Afrikaans). Obviously, however, the varieties with gender marking could further be categorised into those distinguishing two genders (Danish, Dutch, West Frisian) and those distinguishing three genders (Icelandic, German). We could dissolve this one complex variable into several single binary variables, but that would distort the relative weight of the variables in relation to each other. Variables with more possible variants would have a greater overall influence on the results, and that's not what we want: We want all variables to contribute evenly to the end result. Alternatively, we could keep those complex variables as one single feature, but describe it with a scale – but those scales would be arbitrary. A lot of dialectometrical studies showed that in the end it is more important to have reasonable variables than to get too obsessed with scales, so we kept it simple – especially considering that we regard what we are doing as just a first case study.

At the end of the day, our analysis is to some extent wrong – but it is not necessarily **completely** wrong. Errors can be of systematic and unsystematic nature: The unsystematic error of wrong or overly simplified facts in single features is tackled by the quantitative nature of the study. While a single data point is in principle unreliable, a larger number of data points greatly improves the overall reliability of the data set: A single data point might be wrong, but it is highly improbable that all of them are wrong, and it is even more improbable that they are all wrong in the same way. However, all of them might be subject to systematic error. To control for this, multiple options are available. The first one is to try to minimise the probability for a systematic error by eradicating the cause, i.e. a single perspective on the phenomenon. Checking whether the data point is supported by different authors and/or cross-theoretically is the best option. However, this is not always feasible – especially when it comes to lesser-researched varieties. If one cannot minimise the possibility of a systematic error, the second best alternative is to keep it as constant as possible. Thus, ironically, the alternative is to use the same source for all varieties. Again, this is seldom feasible, as there is not a whole lot of research that tracks features through all of our chosen 28 varieties.

### *2.4 Quantitative techniques*

Obviously, the first step towards assessing a typological overview is extracting all the relevant data from grammatical descriptions of the varieties and compiling them into one data set. For this, we opted for a simple feature matrix (variety × feature = value) in the form of a spreadsheet. Values are in most cases coded binarily, for example 1 if a feature is attested, 0 if it is absent, or (arbitrarily) 0 for one option and 1 for the other. In some cases, as explained in subsection 2.3, we allowed for an intermediate value of 0.5.

Based on this feature matrix, we first computed a so-called *distance matrix* (Manhattan distance,[5] for reasons of a) the high number of dimensions and b) the binary nature of our data) for all the varieties.[6] This means that for every pair of varieties, the differences in all the values of the feature matrix are taken into account. The point of this is to arrive at a matrix that quantifies the overall differences between each of the varieties, such as Figure 1.

*Table 3: Distance matrix (excerpt).*

| | Afrikaans | Elvdalian | Danish | German | English | North Frisian | Sater Frisian | West Frisian | Faroese | Icelandic | Yiddish |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Elvdalian | 21.538462 | | | | | | | | | | |
| Danish | 15.428571 | 12.878049 | | | | | | | | | |
| German | 14.857143 | 19.317073 | 12.765957 | | | | | | | | |
| English | 20.923077 | 27.692308 | 21.272727 | 25.090909 | | | | | | | |
| North Frisian | 14.000000 | 17.333333 | 10.536585 | 11.121951 | 19.800000 | | | | | | |
| Sater Frisian | 14.270270 | 16.200000 | 13.463415 | 10.536585 | 21.538462 | 3.243243 | | | | | |
| West Frisian | 14.634146 | 16.000000 | 12.800000 | 11.200000 | 22.285714 | 2.400000 | 3.076923 | | | | |
| Faroese | 22.243902 | 15.000000 | 10.604651 | 12.279070 | 26.400000 | 14.918919 | 13.894737 | 14.285714 | | | |
| Icelandic | 25.800000 | 15.384615 | 12.545455 | 14.181818 | 24.585366 | 17.052632 | 17.684211 | 14.285714 | 5.853659 | | |
| Yiddish | 12.292683 | 19.200000 | 17.066667 | 9.066667 | 22.883721 | 13.200000 | 11.400000 | 13.090909 | 15.428571 | 21.142857 | |
| Limburgish | 15.567568 | 20.108108 | 17.846154 | 8.615385 | 21.333333 | 12.000000 | 12.000000 | 9.729730 | 17.684211 | 15.157895 | 13.621622 |

Based on this distance matrix, we were able to perform two kinds of analyses, a hierarchical cluster analysis and multidimensional scaling. Both techniques are based on the same distance matrix – however, they tend to yield slightly different results, because they treat variation differently and are performed to different ends.

Cluster analysis is done to find groups in the data. These groups should ideally be a) as homogeneous as possible internally while at the same time b) as different from each other as possible. Hierarchical cluster analysis progresses stepwise. First, all varieties are treated as their own, small, separate cluster. Then, an algorithm combines (or fuses) the two clusters that it deems most similar – those two now form a new cluster. Following that, it re-evaluates how similar those new clusters are to each other and again fuses the closest two. This goes on until everything is merged into one big cluster. The researcher is left to decide which number of clusters is the ideal grouping of the data.

Multidimensional scaling, on the other hand, does not yield groups of data, and it does not progress on a step-by-step basis. Its purpose is to reduce higher numbers of dimensions – as are present in our data – to a lower, more easily interpretable number. This is aided by the observation that a lot of the dimensions present in typical variationist data are rather similar. Technically, a MDS attempts to find a mapping from its high-dimensional input to a model that preserves the distances between the original data points in lesser dimensions. It could be compared to trying to take a (two-dimensional) photograph of a (three-dimensional) room from the best angle possible, with the intention that the photograph roughly captures how the objects in the room are placed with respect to each other – although in practice, the numbers of dimensions involved are usually a lot higher.

[5] Manhattan distance is the sum of the absolute distances of the Cartesian coordinates, like manoeuvring a chequerboard (as opposed to i.e. Euclidean distance, which measures distance “as the crow flies”).
[6] For all of the computational steps taken in this study we used the base distribution of the *R* environment, with the addition of the package *ape* (Paradis & Schliep 2019), which was used for visualising the cladograms of Figures 2 to 5.

# 3 Results

## *3.1 Cluster analysis*

The results of a cluster analysis can be visualised using a cladogram, which is a type of tree structure. Figure 2 displays such a structure, based on the clustering of our distance matrix using the Ward algorithm. The tree is built up from the right to the left, with each node representing a cluster. The terminal nodes on the right represent the smallest possible clusters, i.e. the individual varieties. As we move to the left, the clusters get bigger (meaning they contain increasingly more varieties) while becoming fewer. The length of the branches indicates the distance between clusters, i.e. the expense of fusing them into one.

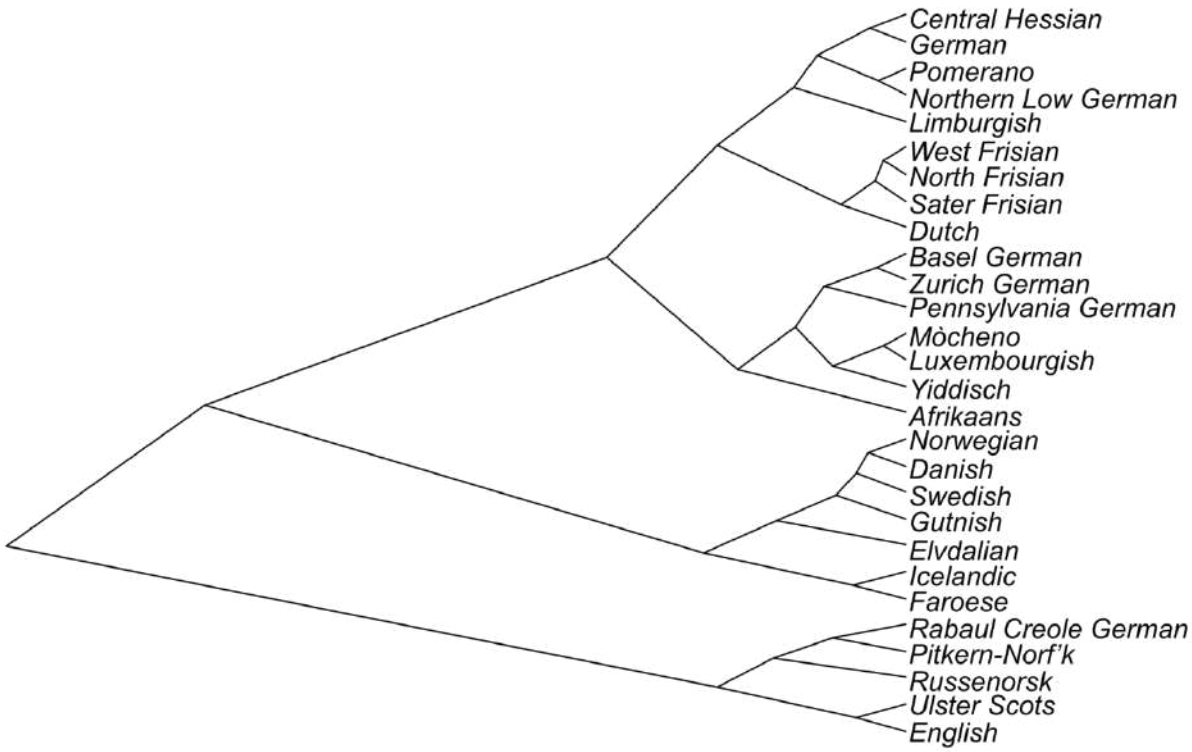


*Figure 2: Cladogram of a hierarchical cluster analysis (Ward algorithm, Manhattan distance).*

In discussing the trees, however, we will be looking at them from left to right. That is, we will start by looking at the large clusters first and then see how they break down into smaller ones. (It is important to note that, while such tree structures are also known in *tree model* or *Stammbaum* type models of language change, they should not automatically be interpreted as any kind of proposition on the diachrony of the varieties involved – certainly not in our case, where relations are to be understood as purely synchronic.)

The respective length of the branches between nodes suggests that cutting the tree at three clusters is a reasonable first choice (Figure 3). In doing so, we arrive at what can be called a Continental West Germanic cluster (at the top, in ocher), a North Germanic cluster (in light green), and a third cluster (in red) containing English, the closely related Ulster Scots, and the pidgins/creoles Pitkern-Norf'k, Russenorsk and Rabaul Creole German (Unserdeutsch). This shows that English (as well as its close relative Ulster Scots) is rather atypical. This is partly due to the fact that about a fifth of our features goes back to Mc Whorter (2002, 2004), who chose them precisely to show that English is different from the other languages (see section

2.2). Nonetheless, there are the four other fifths of our features. Also, the finding is in line with the previous literature more generally, and apparently, what works for English also works for the pidgins/creoles in our sample. So the features selected are suitable for more than just a differentiation of [± English].[7]

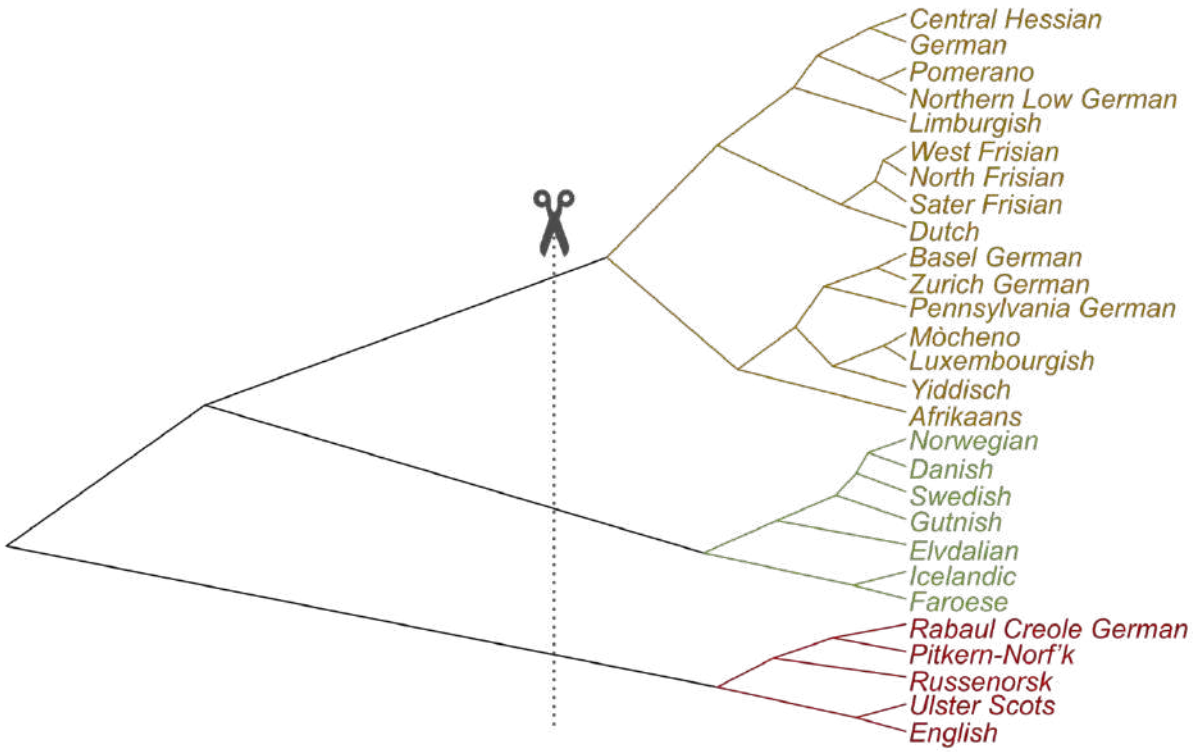


*Figure 3: Cladogram, cut at three clusters.*

The fact that we arrive at a stable North Germanic cluster can be seen as lending a certain degree of validity to our approach as well: Even though North Germanic is decidedly not the focus of our feature selection, the results are linguistically highly plausible. In general, hierarchical clustering can be susceptible to changes of the algorithm and/or the method of computing a distance matrix, so it is important to stress that this North Germanic cluster is totally immune to such changes, as attested in our respective test runs: the North Germanic cluster remains consistently obvious in all other combinations of complete linkage, Ward, Manhattan distance or Euclidean distance and yields significant stability in bootstrapping approaches. Concerning Afrikaans, it is interesting to note that it is not part of our outlier cluster (English, pidgins/creoles) but part of the Continental West Germanic cluster.

If we cut the tree at four clusters (Figure 4), we see that the large West Germanic group splits in two: The result is one group (in grey) that includes Dutch, German, Limburgish, Northern Low German, Pomerano, Central Hessian, and the three Frisian varieties, and another group (in ocher) that includes Zurich German, Basel German, Pennsylvania German, Mòcheno, Luxembourgish, Yiddish and Afrikaans. The first group appears to be more homogeneous in certain respects: It almost exclusively contains varieties that are (mainly) spoken in

[7] In this context, it is interesting to recall Bailey & Maroldt's (1977) hypothesis that Middle English was a creole resulting from the contact with Norman French. See, however, Thomas & Kaufman (1988) for a rebuttal of this thesis.

continental Europe, such as the big standard languages Dutch and German, as well as most of the smaller languages/varieties, with the exception of the four southernmost ones (Luxembourgish, Basel German, Zurich German and Mòcheno). The only variety spoken outside of Europe that is contained in the cluster is Pomerano, a daughter language of (East) Low German. Given its genealogical descent, its closeness to the northern continental varieties still appears plausible.

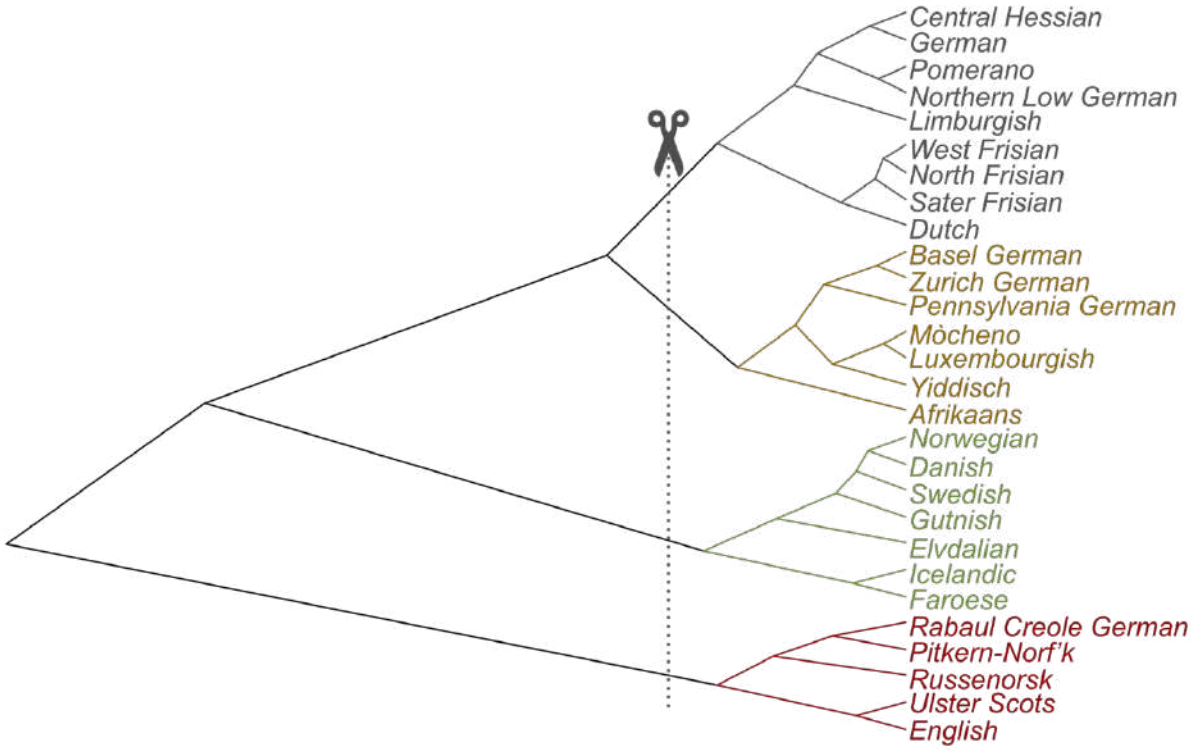


*Figure 4: Cladogram, cut at four clusters.*

The other group appears to be more heterogeneous: With the exception of Pomerano, it contains all the Continental West Germanic varieties that are spoken outside of Europe – Pennsylvania German, Afrikaans – as well as Yiddish, which is spoken worldwide. It also contains all of the varieties known for a high degree of language contact (including, apart from the varieties just mentioned, also Mòcheno) that are not pidgins or creoles. But it also includes the southern varieties of continental Europe (Luxembourgish, Zurich German, Basel German, Mòcheno), most of which would probably not be considered typical contact varieties (Luxembourgish, Zurich German, Basel German).

The high structural similarity of Mòcheno and Luxembourgish might surprise. It is probably due to the combination of two distinct factors. One is the type of contact situation with their respective Romance neighbours, the other one is the relative sparseness of definitive, reliable information on some of our features in the case of Luxembourgish, resulting in more NAs in our data matrix than for most other varieties. It is quite possible that these missing data points might contain information that would further distinguish Luxembourgish from Mòcheno.[8]

[8] It should be noted, however, that the availability of data (probably not entirely coincidentally down to the individual feature) on Limburgish is similarly problematic – yet, they do not form any kind of "bad data"-cluster together, but are conclusively distinct, despite their geographical proximity.

As far as Afrikaans is concerned, it is interesting to note that it does not pattern with the group that includes Dutch, but with the second group, which includes other non-European and high-contact varieties. We commented above that the North Germanic cluster is robust even under changes of the underlying math, i.e., different clustering algorithms or methods to determine variety distances. This is not the case for Afrikaans. Depending on the methods chosen, the four-cluster-solution alternatively groups Afrikaans with the outlier-group (English et al.).[9] Obviously, it never clusters with North Germanic, but, interestingly, it also does not ever cluster with the group containing Dutch, either. In any case, this suggests that Afrikaans is relatively isolated in terms of structure and does not share too many properties with any of the larger groups. We will have to revisit this notion later on.

The five, six and seven cluster solutions are extremely close, so it would not make much sense to further differentiate between them. The seven cluster solution (Figure 5) would show the following changes: Firstly, it would break up the North Germanic cluster, showing a division between insular Scandinavian (Icelandic, Faroese) on the one hand and the varieties spoken in Mainland Scandinavia on the other (Norwegian, Danish, Swedish, Gutnish, Elvdalian).

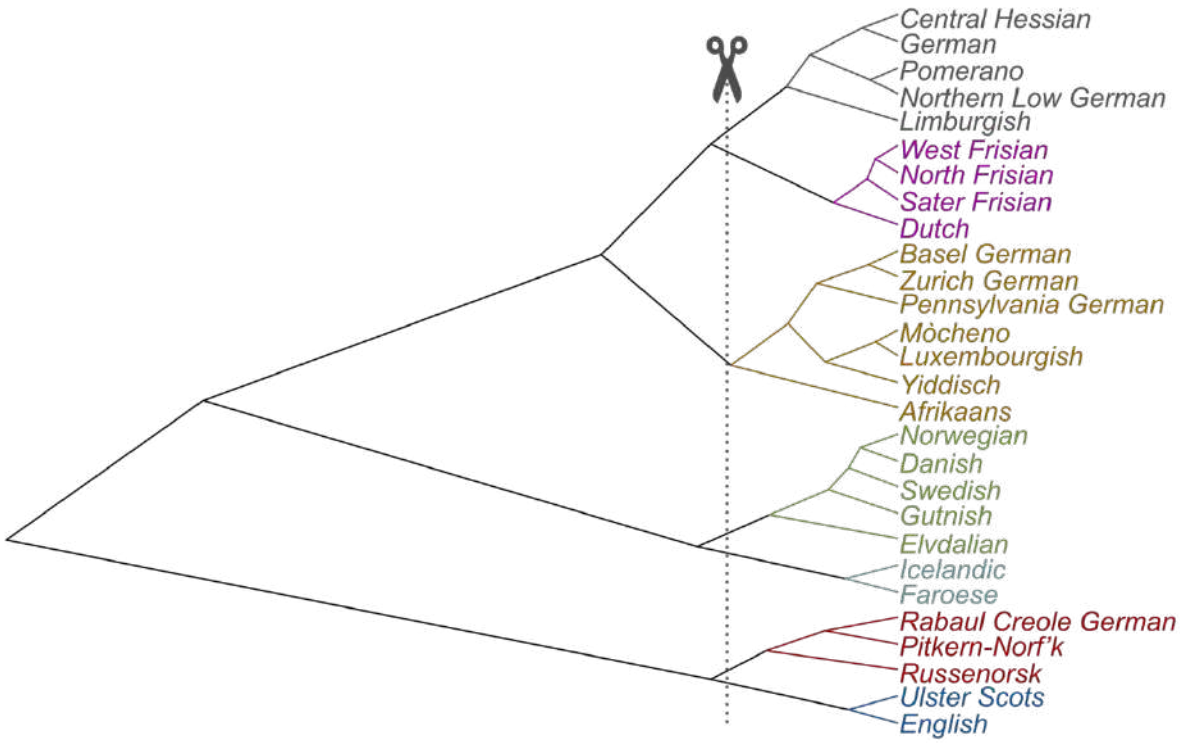


*Figure 5: Cladogram, cut at seven clusters.*

Judging from the existing literature, this is largely to be expected – except, perhaps, for the position of Elvdalian, which might be expected to structurally belong to the insular Scandinavian branch. In any case, Elvdalian is the one that is most different from the rest of

[9] To be more precise, this happens if we compute the distance matrix using Euclidean instead of Manhattan distance. The combination of Manhattan distance and the Ward algorithm – that our analyses are based on – can generally be seen as the mathematically and philologically most reasonable approach to our specific data, so the use of Euclidean distance would constitute a slightly "worse" measure anyway. Nonetheless, as the comparison of these different numerical approaches sheds an interesting light on the volatile classification of Afrikaans, we chose to include it.

the continental varieties. There has been considerable language contact between the continental varieties in the past, but it comes as no surprise that Danish and Norwegian, where this contact has arguably been the strongest, turn out to be the closest to each other. Please bear in mind that this might probably not apply to the spectrum of Danish and Norwegian varieties as a whole – we would expect that, for example, traditional Vestjysk in Denmark and Nordvestlandsk in Norway show considerably more typological differences between each other than those captured by the standard or semi-standard grammars (focusing on Urban East Norwegian and the Copenhagen Standard) that form the basis of our sample.

Secondly, the first sub-group within Continental West Germanic breaks in two. This results in a group with High and Low German, Limburgish, Central Hessian and Pomerano on the one hand, and a group containing Dutch and the three varieties of Frisian on the other. Additionally, the cluster of "outliers" branches into an English/Scots-subgroup and a pidgin/creole-group containing Rabaul Creole German, Pitkern-Norf`k and Russenorsk.

Interestingly, the first varieties that would emerge as their own exclusive clusters would be Afrikaans (in an eight cluster solution) and Elvdalian (nine clusters). This seems to indicate that, with respect to our features, Afrikaans is an outlier, but not in the same way as for example English is. It does not emerge as a single variety in contrast to major subgroups (such as our Continental West Germanic group, our North Germanic group and our 'outlier group'). Nor does it form part of the outlier group, where some might expect it to be located due to its history of high contact. Rather, Afrikaans clusters with the other Continental West Germanic languages. Within that group, however, it appears to be the most isolated one, and it is more similar to a lot of smaller high-contact varieties than it is to Dutch.[10]

### *3.2 Multidimensional scaling*

In a second step, we may now use the same distance matrix that we used for the cluster analyses (3.1) to conduct *multidimensional scaling* (or short *MDS*) of the varieties. While cluster analysis creates discrete groups, multidimensional scaling helps us to detect continuous, flowing relationships in the data. Ideally, the results from both methods should not contradict but complement each other, by providing two different perspectives on the same object. Figure 6 is a visualisation of a MDS that condenses the variation in the distance matrix to just two dimensions (of course having the advantage that two dimensions can be neatly presented on a flat piece of paper – bear in mind that the dimensions themselves have no inherent meaning). The varieties are organised spatially in a way that reflects their differences in the underlying higher-dimensional dataset.

[10] On a slightly smaller scale, something similar can be said about Elvdalian. It is without doubt part of the structural bounds of the North Germanic varieties, but among the continental ones, it is relatively distant to the large standardised national varieties (That is, if we might call Gutnish "continental" despite being spoken on an island. The arguments pro or contra are slightly circular and need not bother us here, but there is reason to treat Gutnish as a branch of the continental varieties on historical/genealogical grounds.) Elvdalian is thus both an outlier among the continental North Germanic varieties and it is considerably different from the insular varieties. Certainly, the typological evidence does not suggest that it is a daughter language or dialect of (Standard) Swedish in a synchronous sense.

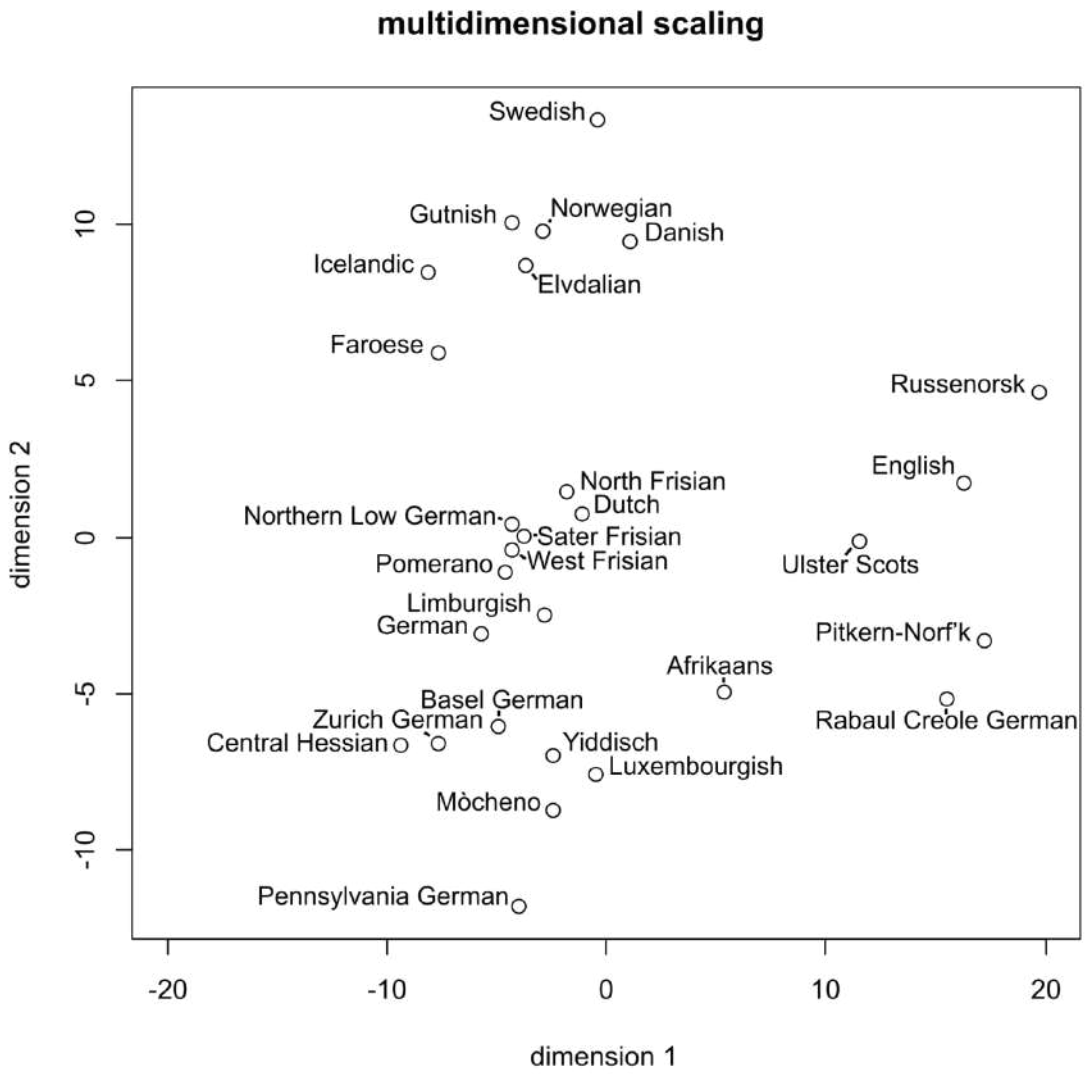


*Figure 6: Multidimensional scaling of the distance matrix of all 28 varieties (goodness of fit ≈ 0.507).*

As could already be seen in the cluster analysis, and as highlighted in Figure 7, the varieties can fairly neatly be divided into two lumps that from a historical perspective could be labelled as North Germanic (at the top) and West Germanic (at the bottom), and a third group containing English/Scots and the pidgins/creoles (to the right). Again, also visible is a division of the North Germanic varieties into insular and continental.

Focusing on West Germanic, the historically and spatially closely connected standard and non-standard varieties of continental Europe turn out to be grouped nicely together. Closest to each other are Northern Low German, Saterland Frisian, West Frisian and Pomerano.

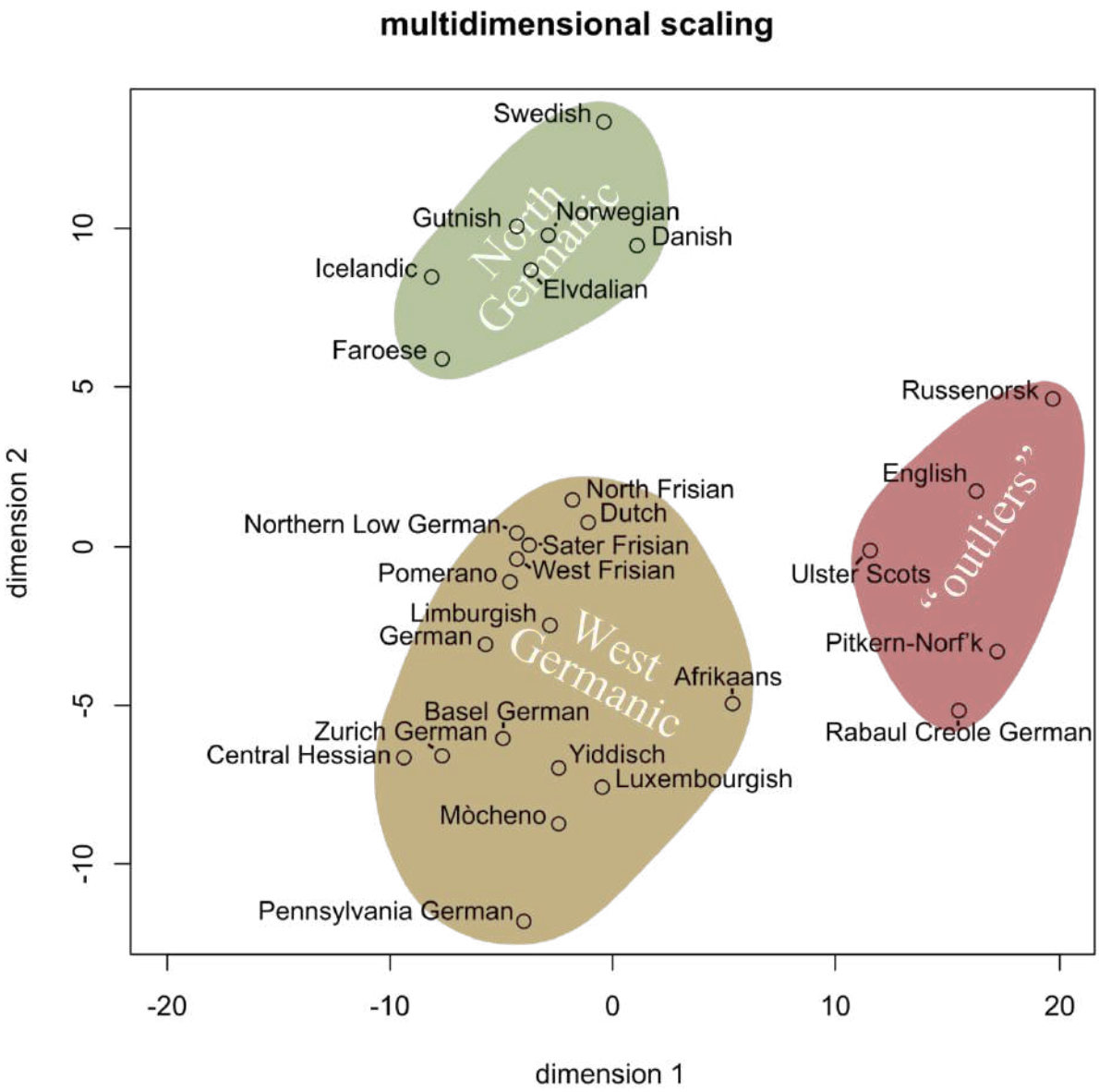


*Figure 7: Multidimensional scaling of the distance matrix of all 28 varieties (historical families).*

By contrast, Afrikaans, as well as Pennsylvania German, are at the periphery of the West Germanic varieties. Afrikaans is functionally even further from its mother language Dutch than from the Southern Bavarian language island of Mòcheno. Synchronically, this is a strong argument against subsuming Afrikaans as a mere "variety of Dutch". In fact, according to our metric, the structural difference between Dutch and Afrikaans is greater than that between Dutch and Faroese or English and Russenorsk. But at the same time, while showing similar tendencies in grammatical reduction, Afrikaans is decisively different from English as well, so contact and reduction of grammatical complexity in Germanic languages need not lead to the same results.

### *3.3 Combined approaches*

As stated in section 2.4, clustering and multidimensional analysis should ideally complement each other instead of providing contradictory evidence. Because of that, we cross-validate both approaches by combining them (as done e.g. in Pröll, Elspaß & Pickl 2021). For Figure 9,

the varieties in the multidimensional scaling plot are coloured according to the clustering results in Figure 5 (section 3.1). It is obvious that the results are at large consistent with each other: Varieties that clustered together also form a coherent spatial pattern in MDS.

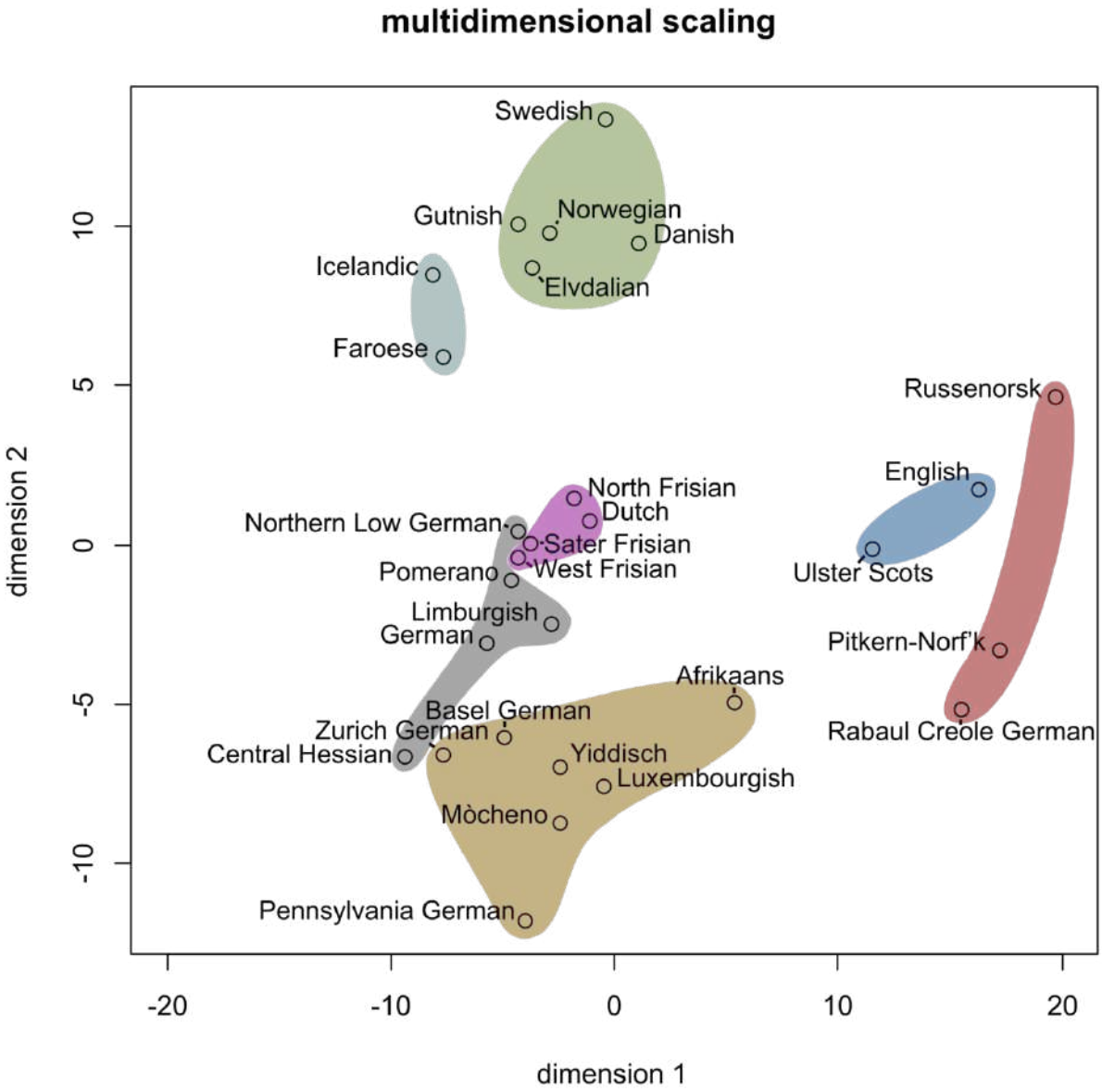


*Figure 9: Combination of MDS (space) and cluster analysis (colour); 7 clusters.*

Afrikaans may be worth a second look: Its middle position between the West Germanic and the pidgin/creole languages is clearer in the MDS than it is in the cluster results. But, as already implied in the passage dealing with the clustering results (section 3.1), the attribution to the West Germanic cluster is less robust for Afrikaans than it is for a lot of the other varieties. Depending on the kind of measure for the distance matrix and the choice of clustering algorithm, it might as well cluster (loosely) with the "outlier"-group. There have been discussions in the past about the origins of Afrikaans, with one of the central controversies surrounding the question of whether it is creolized or an iteration of Dutch (see e.g. Roberge 2010 for an overview). The quantitative structural approach seems to indicate that it is neither to full extent.

From a more general perspective, it is also interesting to note what patterns are conspicuously absent from both the clustering and the MDS results. Most importantly, there is no obvious pattern that implies the degree of standardisation as a clear factor.

## 4 Conclusion

In this article, we aimed to take a look at the typological position of Afrikaans among its Germanic sisters from a quantitative perspective, based on phonological, morphological and syntactic features, including both standard and non-standard varieties as well as other high-contact varieties. As a primary result, our metric yields an expectedly strong divide between the two branches historically labelled as North Germanic and West Germanic, even if we focus on structural/typological features. It is thus synchronically based on far more than just lexical differences or bifurcation in sound change.

While Afrikaans unsurprisingly turns out to be part of the West Germanic branch, it is neither a very "typical" West Germanic variety nor is there synchronic evidence of Afrikaans being a daughter language of Dutch – on the contrary, when seen through a functional/typological perspective (as opposed to, say, one that focuses on lexis or on pronunciation, such as Heeringa / de Wet 2008), Afrikaans appears to be a highly autonomous variety at the structural periphery of West Germanic. But the analysis also shows a group of distinct "outliers", including English: contrary to the preconception of English being a North/West Germanic hybrid, it does not turn out to be a mere mixture of features from both branches. Rather, it is part of a distinctive third branch of varieties, the ones with a specific kind of language contact. Afrikaans, also having a history of intense contact, may be the most similar to those "outliers", but it cannot easily be qualified as one of them.

The analyses suggest that there are two different kinds of "high contact" varieties: On one side, some might have been in close contact with other varieties but are (and have been for centuries) almost exclusively acquired and passed on as a first language by their speakers (such as Pennsylvania Dutch, Mòcheno, Yiddish, Luxembourgish). Their contact languages serve as second languages or parallel first languages in those cases. On the other side, there are "high contact" varieties that somewhere along their history went through a bottleneck situation in acquisition, where a significant amount of speakers came from a different first language and acquired (or even constructed) the variety in question as a second language (most notably Russenorsk, Rabaul Creole German and Pitkern-Norf'k, but also to some extent Afrikaans and English).

On a methodological note, our approach required a number of decisions that might well have been made in a different way – most crucially perhaps the selection and the conceptualisation of the structural features. It would thus be presumptuous to regard our classification of Germanic standard and non-standard varieties as even close to definitive. Nonetheless, the mixture of the results' plausibility (concerning relations on which assumptions and case studies have been drawn before) and new insights (where structural (dis-)similarities as of now had been largely uncharted) suggests that a combined quantitative/functional approach can be feasible and fruitful.

## 5 References


Allan, Robin, Philip Holmes, and Tom Lundskær. 1995. *Danish: A Comprehensive Grammar.* London: Routledge.

Bailey, Charles-James, and Karl Maroldt. 1977. "The French lineage of English." In *Langues en contact – Pidgins – Creoles – Languages in contact*, edited by Jürgen Meisel, 21–53. Tübingen: Narr.

Braunmüller. Kurt. 2000. "Gender in North Germanic: A Diasystematic and Functional Approach." In *Gender in Grammar and Cognition*, edited by Barbara Unterbeck, Matti Rissanen , Terttu Nevalainen, and Mirja Saari, 25–54. Berlin: De Gruyter Mouton.

Chambers, J.K. 2004. "Dynamic typology and vernacular universals." In *Dialectology Meets Typology*, edited by Bernd Kortmann, 127–145. Berlin, New York: De Gruyter, Mouton.

Chomsky, Noam. 1995. "Language and nature." *Mind* 104 (413): 1–61.

Dahl, Östen. 2000. "The grammar of future time reference in European languages." In *Tense and aspect in the languages of Europe*, edited by Östen Dahl, 309–328. Berlin, New York: De Gruyter.

Deumert, Ana, and Wim Vandenbussche. 2003. *Germanic Standardizations: Past to Present*. Amsterdam, Philadelphia: Benjamins. https://doi.org/10.1075/impact.18.

Donaldson, Bruce C. 1993. *A Grammar of Afrikaans.* Berlin, New York: De Gruyter.

Dryer, Matthew S. 2005. "Order of object and verb." In *The World Atlas of Language Structures*, edited by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 338–341. Oxford: Oxford University Press.

Dützmann, Heinz. 1939. "Syntax von Nomen und Verb im Ostlüneburgischen auf Grund der Mundart von Kaarßen." *Zeitschrift für Mundartforschung* 15 (1): 1–24.

Elmentaler, Michael, and Felix Borchert. 2012. "Niederdeutsche Syntax im Spannungsfeld von Kodex und Sprachpraxis." In *Niederdeutsche Syntax*, edited by Robert Langhanke, Kristian Berg, Michael Elmentaler, and Jörg Peters, 101–136. Hildesheim, Zürich, New York: Olms.

Fleischer, Jürg. 2015. "Pro-Drop und Pronominalenklise in den Dialekten des Deutschen." In Deutsche Dialekte: Konzepte, Probleme, Handlungsfelder. Akten des 4. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD), edited by Michael Elmentaler, Markus Hundt, and Jürgen Erich Schmidt, 191–210. Stuttgart: Steiner.

Haider, Hubert. 2013. *Symmetry breaking in syntax*. Cambridge: Cambridge University Press.

Harbert, Wayne. 2007. *The Germanic Languages.* Cambridge: Cambridge University Press.

Hasselberg, Joachim. 1979. Differenzgrammatik Mittelhessisch: Hochsprache. Eine Untersuchung dialektspezifischer Kommunikationsbehinderungen von hessischen Schülern. Gießen: Schmitz.

Haugen, Einar. 1966. "Dialect, language, nation." American Anthropologist 68: 922–935.

Heeringa, Wilbert, and Febe de Wet. 2008. "The origin of Afrikaans pronunciation: a comparison to west Germanic languages and Dutch dialects." In *Proceedings of the 19th Annual Symposium of the Pattern Recognition Association of South Africa*, edited by Fred Nicolls, 159–164. Cape Town: Pattern Recognition Association of South Africa.

Jahr, Ernst Håkon. 2005. "Russenorsk. Language contact in Scandinavia in the 19th century." In *The Nordic languages: An international handbook of the history of the north Germanic languages. Volume 2*, edited by Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman, 1538–1542. Berlin, New York: De Gruyter.

Kehrein, Roland. 2009. "Dialektalität von Vorleseaussprache im diatopischen Vergleich. Hörerurteil und phonetische Messung." *Zeitschrift für Dialektologie und Linguistik* 76 (1): 14–54.

Kortmann, Bernd. 2004. *Dialectology Meets Typology*. Berlin, New York: Mouton de Gruyter.

Kortmann, Bernd, Kerstin Lunkenheimer, and Katharina Ehret. 2020. *The Electronic World Atlas of Varieties of English.* Zenodo. Accessed August 19 2023. https://doi.org/10.5281/zenodo.3712132. http://ewave-atlas.org.

Kramer, Pyt. 1982. *Kute Seelter Sproakleere.* Rhauderfehn: Ostendorp Verlag.
Lameli, Alfred. 2013. Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland. Berlin: De Gruyter.
Lindow, Wolfgang, Dieter Möhn, Hermann Niebaum, Dieter Stellmacher, Hans Taubken, and Jan Wirrer. 1998. *Niederdeutsche Grammatik*. Leer: Schuster.
McWhorter, John. 2002. "What happened to English?" *Diachronica* 19: 217–272.
McWhorter, John. 2004. "What happened to English?" In *Focus on Germanic Typology*, edited by Werner Abraham, 19–60. Berlin: Akademie.
Murelli, Adriano, and Bernd Kortmann. 2011. "Non-standard varieties in the areal typology of Europe." In *The Languages and Linguistics of Europe: A Comprehensive Guide*, edited by Bernd Kortmann, and Johan van der Auwera, 525–544. Berlin, New York: De Gruyter, Mouton.
Paradis, Emmanuel, and Klaus Schliep. 2019. "ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R." *Bioinformatics* 35: 526–528. doi:10.1093/bioinformatics/bty633.
Postma, Gertjan. 2014. "Complex complementizers and the structural relation with weak T. New (morpho)syntactic data from a Pomeranian language island in Brazil." *Stuf* 67(4): 627–663.
Pröll, Simon, Stephan Elspaß, and Simon Pickl. 2021. "Areal microvariation in German-speaking urban areas (Ruhr Area, Berlin, and Vienna)." In *Urban Matters. Current Approaches of International Sociolinguistic Research*, edited by Arne Ziegler, Stefanie Edler, Nina Kleczkowski, and Georg Oberdorfer, 227–252. Amsterdam, Philadelphia: Benjamins.
Roberge, Paul T. 2010. "Afrikaans: considering origins." In *Language in South Africa*, edited by Rajend Mesthrie, 79–103. Cambridge: Cambridge University Press.
Seiler, Guido. 2019. "Non-Standard Average European." In *‚athe in palice, athe in anderu sumeuuelicheru stedi': Raum und Sprache: Festschrift für Elvira Glaser zum 65. Geburtstag*, edited by Andreas Nievergelt, and Ludwig Rübekeil, 541–554. Heidelberg: Winter.
Seiler, Guido, Oliver Schallert, and Thilo Weber. under review. Submitted to Joshua Bousquette & Simon Pickl (eds.), *Oxford Handbook of the German Language*. Oxford: Oxford University Press.
Smits, Tom. 2011. Strukturwandel in Grenzdialekten. Die Konsolidierung der niederländisch-deutschen Staatsgrenze als Dialektgrenze. Stuttgart: Steiner.
Thieroff, Rolf. 2000. "On the areal distribution of tense-aspect categories in Europe." In *Tense and aspect in the languages of Europe* , edited by Östen Dahl, 264–305. Berlin: De Gruyter.
Thomasson, Sarah Grey, and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.
Trudgill, Peter. 2011. Sociolinguistic Typology. Social Determinants of Linguistic Complexity. Oxford: Oxford University Press.
Weber, Albert. 1948. *Zürichdeutsche Grammatik.* Zürich: Schweizer Spiegel Verlag.
Åkerberg, Bengt. 2012. Älvdalsk grammatik. Under medverkan av Gunnar Nyström. Mora: Ulum Dalska.

# 6 Appendix A: Illustration of our features

## *6.1 Gender marking on the definite article*

e.g. Dutch ***de*** *man* (the.COMMON man 'the man') vs. ***het*** *kind* (the.NEUTER child 'the child')

## *6.2 Case marking on the definite article*

e.g. German ***der*** *Tisch* (the.NOM table 'the table') vs. ***den*** *Tisch* (the.ACC table 'the table')

## *6.3 (Loss of) derivational prefixes*

loss of forms corresponding to e.g. Afrikaans *be-* (as in ***be****slis* 'decide'), *ont-* (as in ***ont****ken* 'deny') or ***ver-*** (as in *verpletter* 'to smash') (cf. McWhorter 2004: 31—33).

## *6.4 Number distinction in second person pronouns*

e.g. Swedish *du* (you.SING) vs. *ni* (you.PL)

## *6.5 Dual category in pronouns*

e.g. Elvdalian *ðieɾ fingg* ***weɾ*** *saina* (they got each.PL theirs 'each got theirs') vs. *ðieɾ åvå faið* ***wåɾ*** *saina* (they have gotten each.DUAL theirs 'both got theirs') (cf. Åkerberg 2012: 241)

## *6.6 Subject-verb-agreement*

e.g. English *I live* vs. *she live****s***

## *6.7 Syncretism of verbal singular forms*

e.g. Dutch (2nd person = 3rd person): *jij* ***werkt*** (you.SG work), *zij* ***werkt*** (she works)

## *6.8 Syncretism of verbal plural forms*

e.g. Northern Low German (1st person = 2nd person = 3rd person): *wi* ***loopt*** (we run), *ji* ***loopt*** (you.PL run), *se* ***loopt*** (they run)

## *6.9 'Ge'-prefix on past participles in perfect constructions*

e.g. *Dutch Ik heb* ***ge****werkt* 'I have worked'

## *6.10 Past subjunctive forms (distinct from past indicative) beyond 'be'*

e.g. German *Ich* ***kam*** (I come.PST.IND) vs. *Ich* ***käme*** (I come.PST.SUBJ)

## *6.11 Loss of the synthetic preterite*

loss of forms corresponding to e.g. English *she* ***came***, *she* ***lived***

## *6.12 New synthetic passive voice*

e.g. Danish *Posten udbringe****s*** *hver dag* (post-the deliver-S every day 'The post is delivered every day', Allan et al. (1995: 319), quoted from Harbert (2007: 328)

### *6.13 Presence of an infinitive suffix*

e.g. Dutch *-en* as in *lop**en*** (run-INF ‘to run’)

### *6.14 (Reflex of the) distinction between strong and weak adjective inflection*

e.g. Danish *en stor_ bil* (‘a big car’) vs. den *stor**e** bil* (‘the big car’), adapted from Allan et al. (2000: 35)

### *6.15 Subject-agreement marking on predicative adjectives*

e.g. Danish *bilen er stor_* (car-the.COMMON is big) vs. huset er stor**t** (house-the.NEUTER is big.NEUTER), adapted from Allan et al. (2000: 35)

### *6.16 Morphological number marking on nouns*

e.g. English *two dog**s*** vs. no attestation of an equivalent in Russenorsk (Broch & Jahr 1984)

### *6.17 Pro-drop possible*

e.g. Pennsylvania German *wi bɪʃt* (how be.2SG ‘how are you’), Van Ness (1994: 429)

### *6.18 Pronoun doubling for emphasis*

e.g. Gutnish ***Han** har nuck pänningar **han*** (he has enough coins he ‘he has enough money’), Gustavson 1977: 36)

### *6.19 Verb second*

e.g. Yiddish *dos bux lejənt der man* (the book reads the man ‘it’s the book that the man is reading’), adapted from Jacobs (2005: 224)

### *6.20 Basic order of verb and object*

e.g. Dutch *een appel eten* (an apple eat-INF ‘eat an apple’) vs. English *eat an apple*

### *6.21 Doubly-filled COMP*

e.g. Pomerano *Wätst duu **woufon dat** air boterfat måkt waard?* (know.2SG you from-what that a butter container made is ‘Do you know where a butter container is made of?’), Postma (2018: 124)

### *6.22 Extended use of the present perfect*

e.g. German *Ich bin gestern gekommen* (I am yesterday come ‘I came yesterday’), Harbert (2007: 312)

### *6.23 ‘Be’ as a perfect auxiliary*

e.g. West Frisian *ik **bin** flein* (I am flown ‘I have flown’), Tiersma (1999: 72)

### *6.24 Double perfect*

e.g. Yiddish *Ven ix bin gekumen,* ***hot*** *er* ***gehat ajngešpant*** *di ferd* (when I am come.PTCP has he have.PTCP harness.PTCP the horses 'When I came, he had (already) harnessed the horses'), Kiefer (1994: 141)

### *6.25 Passive marking with 'become'*

e.g. Afrikaans *Die boek* ***word*** *gelees* (the book becomes read.PTCP 'the book is being read'), Donaldson (1993: 257)

### *6.26 Mediopassive from old reflexive*

e.g. Limburgish *Der sal singt* ***sich*** *legt* (the hall sings REFL easily 'this hall has good acoustics'), Cornips (2013: 379)

### *6.27 Inherent reflexivity marking*

e.g. Faroese *ætla sær* (intend-INF REFL 'to intend'), McWhorter (2004: 23)

### *6.28 Highly grammaticalised progressive*

e.g. English What are you doing? – I'm smoking. / *I smoke, Willkowei (2005: 143)

### *6.29 Highly grammaticalised future*

e.g. English *I'****ll do*** *it tomorrow* vs. e.g. *Dutch Dat* ***doe*** *ik morgen* (that do.PRS I tomorrow 'I'll do it tomorrow'), Donaldson (2008: 181)

### *6.30 Modal verb as a future auxiliary*

e.g. Afrikaans *Ik* ***sal*** *dit doen as ek tyd het* (I shall it do-INF if I time have 'I'll do it if I have the time'), Donaldson (1993: 235)

### *6.31 Distinct habitual construction*

e.g. English She **used to** dance

### *6.32 Presence of a definite article*

e.g. English *the* vs. no attestation of an equivalent in Russenorsk (Broch & Jahr 1984)

### *6.33 Presence of an indefinite article*

e.g. German *ein Mann* ('a man') vs. Icelandic *maður* ('(a) man'), Braunmüller (2007: 251)

### *6.34 Position of the definite article*

e.g. German ***das*** *Mädchen* (the girl) vs. Swedish *flicka****n*** (girl.the) 'the girl', Braunmüller (2007: 50)

### *6.35 External possessors*

e.g. West Frisian *Ik stompte* ***my*** *de holle* (I bump.PAST me.DAT the head 'I bumped my head'), McWhorter (2004: 26)

### *6.36 Multiple negation*

e.g. Afrikaans *Niks gebeur nie* (nothing happens NEG 'Nothing is happening'), Donaldson (1993: 402)

### *6.37 Lexical tone*

e.g. Norwegian [1]*tanken* (= toneme 1, tank-the 'the tank') vs. [2]*tanken* (= toneme 2, thought-the 'the thought'

### *6.38 Front rounded vowels*

e.g. /yː/ as in German *kühn* 'bold'

### *6.39 Diphthongs*

e.g. /aʊ/ as in Gutnish *rauk* (= typical limestone rock formation)

### *6.40 Phonological contrast through nasalized vowels*

e.g. Elvdalian [kʉlːɐ] ('girl') vs. [kʉlːɛ̃] (girl.DET 'the girl')

### *6.41 Distinct vowel quantities*

e.g. Saterland Frisian *Bùk* /ʊ/ (= short, 'can') vs. *Buk* /uˑ/ (= half-long, 'belly') vs. *fuul* /uː/ (= long, 'many'), Kramer (1982: 5–6)

### *6.42 Long consonants*

e.g. Icelandic *hatur* [haːtʰʏr] ('hate') vs. *hattur* [haʰtːʏr] 'hat'

### *6.43 Velar nasal*

e.g. /ŋ/ as in English *sing*

### *6.44 Number of series of fricatives*

e.g. German /f/ /s/ ([+ fortis], [- voice]) vs. /v/ /z/ ([- fortis], [+ voice])

### *6.45 Type of contrast between the two series of plosives*

e.g. predominantly [± aspiration] in Faroese vs. predominantly [± voice] in Dutch

### *6.46 Directional adverbs*

e.g. Swedish *här* ('here') vs. *hit* ('to here') vs. *härifrån* ('from here'), Holmes & Hinchliffe (1997: 115 –116), cited from McWhorter (2004: 35)

### *6.47 Inherited /ɣ/*

e.g. Dutch *goed* /ɣuːt/ (‘good’) vs. German *gut* /guːt/ (‘good’)

## 7 Appendix B: Sources


Abraham, Werner. 1989. “Futur-Typologie in den germanischen Sprachen.” In *Tempus – Aspekt – Modus: Die lexikalischen und grammatischen Formen in den germanischen Sprachen,* edited by Werner Abraham and Theo Janssen, 345–390. Berlin, New York: De Gruyter.

Allan, Robin, Philip Holmes, and Tom Lundskær. 1995. *Danish: A Comprehensive Grammar.* London: Routledge.

Allan, Robin, Philip Holmes, and Tom Lundskær-Nielsen. 2000. *Danish: An essential grammar*. London, New York: Routledge.

Andersson, Erik. 1994. “Swedish.” In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 271–312. London, New York: Routledge.

Arfsten, Antje, Anne Pausen-Schwarz, and Lena Terhart. 2022a. *Friesische Gebrauchsgrammatik – Fering.* Bredstedt: Nordfriisk Instituut.

Arfsten, Antje, Anne Pausen-Schwarz, and Lena Terhart. 2022b. *Friesische Gebrauchsgrammatik – Mooringer Frasch.* Bredstedt: Nordfriisk Instituut.

Bacskai-Atkari, Julia. 2022. “German V2 and Doubly Filled Comp in West Germanic.” *The Journal of Comparative Germanic Linguistics* 23: 125–160.

Barba, Preston A., and Albert F. Huffington. 1954. *A Pennsylvania German Grammar.* Allentown, PA: Schlechter.

Barnes, Michael P., and Eivind Weyhe. 1994. Faroese. In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 190–218.. London, New York: Routledge.

Birkenes, Magnus B., and Jürg Fleischer. 2019. “Zentral-, Nord- und Osthessisch.” In *Sprache und Raum: Ein internationales Handbuch der Sprachvariation. Vol. 4: Deutsch*, edited by Joachim Herrgen, and Jürgen Erich Schmidt, 435–478. Berlin: De Gruyter Mouton.

Bonner, Maria. 1995. “Zum Tempussystem des Isländischen.” In *Tense Systems in European Languages II*, edited by Rolf Thieroff, 37–58. Tübingen: Niemeyer.

Braunmüller, Kurt. 1999. *Die skandinavischen Sprachen im Überblick.* Tübingen: Niemeyer.

Braunmüller. Kurt. 2000. “Gender in North Germanic: A Diasystematic and Functional Approach.” In *Gender in Grammar and Cognition*, edited by Barbara Unterbeck, Matti Rissanen , Terttu Nevalainen, and Mirja Saari, 25–54. Berlin: De Gruyter Mouton.

Broch, Ingvild, and Ernst Håkon Jahr. 1984. *Russenorsk: Et pidginspråk i Norge. 2. utgave.* Oslo: Novus.

Buffett, Alice, and Don Laycock. 1988. *Speak Norfolk today.* Norfolk Island: Himii Publishing.

Cornips, Leonie. 2013. “Recent developments in the Limburg dialect region.” In *Language and Space. An international Handbook of Linguistic Variation. Volume 3: Dutch*, edited by Frans Hinskens, and Johan Taeldeman, 378–399. Berlin, Bosten: De Gruyter Mouton.

Dahl, Östen. 1995. “The Tense System of Swedish.” In *Tense Systems in European Languages II*, edited by Rolf Thieroff, 59–68. Tübingen: Niemeyer.

Dahl, Östen. 2000. “The grammar of future time reference in European languages.” In *Tense and aspect in the languages of Europe*, edited by Östen Dahl, 309–328. Berlin, New York: De Gruyter.

Deumert, Ana, and Wim Vandenbussche. 2003. *Germanic Standardizations: Past to Present*. Amsterdam, Philadelphia: Benjamins. https://doi.org/10.1075/impact.18.

De Schutter, Georges. 1994. “Dutch.” In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 439–477.London: Routledge.

De Schutter, Georges, and Ben Hermans. 2013. “The Limburg dialects: Grammatical properties.” In *Language and Space. An international Handbook of Linguistic Variation. Volume 3: Dutch*, edited by Frans Hinskens, and Johan Taeldeman, 356–377. Berlin, Boston: De Gruyter, Mouton.

De Vaere, Hilde. 2021. “Alternating constructions with ditransitive *geben* in present-day German.” *Corpus Linguistics and Linguistic Theory* 17 (1): 73–107.

Doetjes, Gerke, and Renée van Bezooijen. 2007. “Miniporträt Friesisch.” In *EuroComGerm – Die sieben Siebe: Germanische Sprachen lesen lernen*, edited by Britta Hufeisen, and Nicole Marx, 239–253. Aachen: Shaker.
Donaldson, Bruce C. 1993. *A Grammar of Afrikaans.* Berlin, New York: De Gruyter.
Donaldson, Bruce C. 1994. “Afrikaans.” In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 478–504. London: Routledge.
Donaldson, Bruce C. 2008. *Dutch: A comprehensive grammar. 2nd edition.* London, New York: Routledge.
Dryer, Matthew S. 2005. “Order of object and verb.” In *The World Atlas of Language Structures*, edited by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 338–341. Oxford: Oxford University Press.
Dryer, Matthew S. 2013. “Expression of Pronominal Subjects.” In *The World Atlas of Language Structures Online*, edited by Matthew S. Dryer, and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed August 19 2023. http://wals.info/chapter/101.
Ebert, Karen H. 2000. “Progressive Markers in Germanic languages.” In *Tense and Aspect in the Languages of Europe*, edited by Östen Dahl, 605–653. Berlin, New York: De Gruyter.
Eisenberg, Peter. 2013. Grundriss der deutschen Grammatik. Bd. 2: Der Satz. Vierte Ausgabe. Stuttgart, Weimar: Metzler.
Elmentaler, Michael. 2009. “Modernes Niedersächsisch – Dialektwandel im nordniederdeutschen Raum.” In *Low Saxon Dialects across Borders – Niedersächsische Dialekte über Grenzen hinweg*, edited by Alexandra N. Lenz, Charlotte Gooskens, and Siemon Reker, 339–365. Stuttgart: Steiner.
Elmentaler, Michael, and Felix Borchert. 2012. “Niederdeutsche Syntax im Spannungsfeld von Kodex und Sprachpraxis.” In *Niederdeutsche Syntax*, edited by Robert Langhanke, Kristian Berg, Michael Elmentaler, and Jörg Peters, 101–136. Hildesheim, Zürich, New York: Olms.
Fabricius-Hansen, Cathrine. 1994. “Das dänische und norwegische Tempussystem im Vergleich mit dem deutschen.” In *Tense Systems in European Languages*, edited by Rolf Thieroff, and Joachim Ballweg, 49–68. Tübingen: Niemeyer.
Faarlund, Jan Terje, Svein Lie, and Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk.* Oslo: Universitetsforlaget.
Fertig, David. 2020. “Verbal Inflectional Morphology in Germanic.” In *The Cambridge handbook of Germanic linguistics*, edited by Michael T. Putnam, and B. Richard Page, 193–213. Cambridge: Cambridge University Press.
Fiebertshäuser, Hans. 1987. *Das hessische Dialektbuch*. München: Beck.
Fischer, Hanna. 2016. “Präteritum/Perfekt-Distribution.” *SyHD-atlas*. Accessed August 19 2023. https://www.syhd.info/apps/atlas/index.html#praeteritumperfekt-distribution.
Fleischer, Jürg. 2015. “Pro-Drop und Pronominalenklise in den Dialekten des Deutschen.” In Deutsche Dialekte: Konzepte, Probleme, Handlungsfelder. Akten des 4. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD), edited by Michael Elmentaler, Markus Hundt, and Jürgen Erich Schmidt, 191–210. Stuttgart: Steiner.
Garbacz, Piotr, and Janne Bondi Johannessen. 2015. “Övdalian from 1909 to 2009.” In *Studies in Övdalian Morphology and Syntax. New research on a lesser-known Scandinavian language*, edited by Kristine Bentzen, Henrik Rosenkvist, and Janne Bondi Johannessen, 11–46. Amsterdam, Philadelphia: Benjamins.
Gilles, Peter, and Jürgen Trouvain. 2013. “Illustrations of the IPA: Luxembourgish.” *Journal of the International Phonetic Association* 43 (1): 67–74.
Grønnum, Nina. 1998. “Danish.” Journal of the International Phonetic Association 28: 99–105.
Gussenhoven, Carlos, and Flor Aarts. 1999. “The dialect of Maastricht.” *Journal of the International Association* 29 (2): 155–166.
Gustavson, Herbert. 1977. *Gutamålet. Inledning till studium.* Visby: B. Press.

Haberland, Hartmut. 1994. “Danish.” In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 313–348. London, New York: Routledge.
Haider, Hubert. 2013. *Symmetry breaking in syntax*. Cambridge: Cambridge University Press.
Harbert, Wayne. 2007. *The Germanic Languages.* Cambridge: Cambridge University Press.
Hasselberg, Joachim. 1979. Differenzgrammatik Mittelhessisch: Hochsprache. Eine Untersuchung dialektspezifischer Kommunikationsbehinderungen von hessischen Schülern. Gießen: Schmitz.
Hermans, Ben. 2013. “Phonological features of Limburgian dialects.” In *Language and Space. An international Handbook of Linguistic Variation. Volume 3: Dutch*, edited by Frans Hinskens, and Johan Taeldeman, 336–355.Berlin, Boston: De Gruyter Mouton.
Hoekstra, Jarich F. 2001. “Comparative Aspects of Frisian Morphology and Syntax.” In *Handbuch des Friesischen. Handbook of Frisian Studies*, edited by Horst Haider Munske, 775–786. Tübingen: Niemeyer.
Hoekstra, Jarich, and Peter Meijes Tiersma. 1994. “Frisian.” In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 505–531. London: Routledge.
Höder, Steffen. 2010. “Das Lautsystem des Altenwerder Platt. Eine phonetisch-phonologische Bestandsaufnahme.” *Niederdeutsches Wort* 50: 1–27.
Holmes, Philip, and Ian Hinchliffe. 1997. *Swedish: an essential grammar.* London: Routledge.
Hultman, Tor G. 2003. *Svenska Akademiens språklära.* Stockholm: Svenska Akademien.
Jahr, Ernst Håkon. 2005. “Russenorsk. Language contact in Scandinavia in the 19th century.” In *The Nordic languages: An international handbook of the history of the north Germanic languages. Volume 2*, edited by Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann, and Ulf Teleman, 1538–1542. Berlin, New York: De Gruyter.
Jacobs, Neil G. 2005. *Yiddish. A Linguistic Introduction.* Cambridge: Cambridge University Press.
Jacobs, Neil G., Ellen F. Prince, and Johan van der Auwera. 1994. “Yiddish.” In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 388–419. London: Routledge.
Jordan, Caspar. 2011. “Documentation of Gutnish phonetics and phonology.” Master thesis, SOAS University of London.
Kehrein, Roland. 2009. “Dialektalität von Vorleseaussprache im diatopischen Vergleich. Hörerurteil und phonetische Messung.” *Zeitschrift für Dialektologie und Linguistik* 76 (1): 14–54.
Kiefer, Ulrike. 1994. “Die Tempusformen im Jiddischen.” In *Tense Systems in European Languages*, edited by Rolf Thieroff, and Joachim Ballweg, 135–148. Tübingen: Niemeyer.
Kirbik, Aleksandr. 1998. “Does Intragenetic Typology Make Sense?” In *Sprache in Raum und Zeit. Vol. 2: Beiträge zur empirischen Sprachwissenschaft*, edited by Winfried Boeder, Christoph Schroeder, Karl Heinz Wagner, and Wolfgang Wildgen, 61–68. Tübingen: Narr.
Kramer, Pyt. 1982. *Kute Seelter Sproakleere.* Rhauderfehn: Ostendorp Verlag.
Kress, Bruno. 1982. *Isländische Grammatik*. Munich: Hueber.
Lameli, Alfred. 2013. Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland. Berlin: De Gruyter.
Larsson, Ida. 2014. “Double complementizers.” *Nordic Atlas of Language Structures (NALS) Journal* 1. 447–457. Accessed August 19 2023. https://doi.org/10.5617/nals.5413.
Lindow, Wolfgang, Dieter Möhn, Hermann Niebaum, Dieter Stellmacher, Hans Taubken, and Jan Wirrer. 1998. *Niederdeutsche Grammatik*. Leer: Schuster.
Lochtman, Katja. 2019. “Limburgisch.” In *EuroComGerm. Germanische Sprachen lesen lernen. Seltener gelernte germanische Sprachen*, edited by Janet Duke, 151–186. Düren: Shaker.
Lundskær-Nielsen, Tom, and Philip Holmes. 2010. *Danish: A comprehensive grammar (2nd edition).* London, New York: Routledge.
Maitz, Péter, and Siegwalt Lindenfelser. 2018. “Unserdeutsch: Ein (a)typisches Kreol?” *Zeitschrift für Dialektologie und Linguistik* 85 (3): 307–347.
McWhorter, John. 2002. “What happened to English?” *Diachronica* 19: 217–272.

McWhorter, John. 2004. “What happened to English?” In *Focus on Germanic Typology*, edited by Werner Abraham, 19–60. Berlin: Akademie.
Millar, Robert McColl. 2018. *Modern Scots. An Analytical Survey.* Edinburgh: Edinburgh University Press.
Mühlhäusler, Peter. 1996. “German in the Pacific area.” In *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, edited by Stephen A. Wurm, Peter Mühlhäusler, and Darrell T. Tryon, 345–351. Berlin, New York: De Gruyter.
Mühlhäusler, Peter. 2020. Pitkern-Norf’k. The Language of Pitcairn Island and Norfolk Island. Boston, Berlin: Walter de Gruyter.
Peters, Jörg. 2008. “Saterfrisian intonation. An analysis of historical recordings.” *Us Wurk. Tydskrift foar Frisistyk. Journal of Frisian Studies* 57 (3-4): 141–169. Accessed August 19 2023. https://ugp.rug.nl/uswurk/article/view/27359.
Peters, Jörg. 2019. “Saterland Frisian.” *Journal of the International Phonetic Association* 49 (2): 223–230.
Pétursson, Magnús. 1978. Isländisch. Eine Übersicht über die moderne isländische Sprache mit einem kurzen Abriß der Geschichte und Literatur Islands. Hamburg: Buske.
Postma, Gertjan. 2014. “Complex complementizers and the structural relation with weak T. New (morpho)syntactic data from a Pomeranian language island in Brazil.” *Stuf* 67(4): 627–663.
Postma, Gertjan. 2018. *A Contrastive Grammar of Brazilian Pomeranian.* Amsterdam: Meertens Institute.
Quirk, Randolph, Sidney Greenbaum, and Geoffrey Leech. 1985. *A Comprehensive Grammar of the English Language.* London, New York: Longman.
Rieger, Caroline L. 2019. “Luxemburgisch.” In *EuroComGerm. Germanische Sprachen lesen lernen. Seltener gelernte germanische Sprachen*, edited by Janet Duke, 187–224. Düren: Shaker.
Robinson, Christine. 2012. *Modren Scots Grammar. Wirkin Wi Wirds.* Edinburgh: Luath Press Limited.
Robinson, Philip. 2018. Ulster-Scots. A grammar of the traditional written & spoken language. Third edition. Belfast: Ullans Press.
Rosenkvist, Henrik. 2015. “Negative Concord in Four Varieties of Swedish.” *Arkiv for Nordisk Filologi* 130: 139–165.
Rowley, Anthony. 1986. Fersental (Val Fèrsina bei Trient/Oberitalien). Untersuchung einer Sprachinselmundart. Berlin: De Gruyter.
Sapir, Yair. 2006. “Elfdalian, the Vernacular of Övdaln.” In *Proceedings of the 1st Conference on Övdalian, Älvdalen, June 18th–19th, 2004.* Accessed August 19 2023. http://www.diva-portal.org/smash/get/diva2:773379/FULLTEXT02.pdf.
Seiler, Guido. 2004. “On three types of dialect variation and their implications for linguistic theory. Evidence from verb clusters in Swiss German dialects.” In *Dialectology meets Typology. Dialect Grammar from a Cross-Linguistic Perspective*, edited by Bernd Kortmann, 367–399. Berlin, New York: De Gruyter.
Schallert, Oliver, Alexander Dröge, and Jeffrey Pheiff. 2018. *Doubly-filled COMPs in Dutch and German: A bottom-up approach.* Unpublished manuscript, University of Munich and University of Marburg.
Schlitz, Luc. 2012. Lëtzebuergesch schwätzen. Einblick in die luxemburgische Sprache. Trier: Verlag Michael Veyland.
Schmidtkunz, Lena. 2016. “Die Morphosyntax des Verbs in der Kreolsprache Unserdeutsch: Evidenz aus einem Minikorpus.” Master thesis, LMU Munich.
Suter, Rudolf. 1992. *Baseldeutsch-Grammatik. 3., überarbeitete Auflage.* Basel: Christoph-Merian-Verlag.
Thieroff, Rolf. 2000. “On the areal distribution of tense-aspect categories in Europe.” In *Tense and aspect in the languages of Europe* , edited by Östen Dahl, 264–305. Berlin: De Gruyter.

Thieroff, Rolf. 2004. “The subjunctive mood in German and in the Germanic languages.” In *Focus on Germanic Typology*, edited by Werner Abraham, 315–358. Berlin: De Gruyter.
Þráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen, and Zakaris Svabo Hansen. 2004. *Faroese: An overview and reference grammar.* Tórshavn: Føroya Fróðskaparfelag.
Tiersma, Pieter M. 1999. *Frisian Reference Grammar.* Ljouwert: Fryske Akademy.
Van Ness, Silke. 1994. “Pennsylvania German.” In *The Germanic Languages*, edited by Ekkehard König, and Johan van der Auwera, 420–438. London: Routledge.
Volker, Craig Alan. 1982. *An Introduction to Rabaul Creole German (Unserdeutsch).* Master thesis, University of Queensland.
Weber, Albert. 1948. *Zürichdeutsche Grammatik.* Zürich: Schweizer Spiegel Verlag.
Weber-Messerich, Jackie, and Franck Colotte. 2014. *Le Luxembourgeois à grande vitesse.* Köln: Assimil.
Weijenberg, Roger. 2021. *Mestreechter Taol.* Accessed August 19 2023. http://www.mestreechtertaol.nl.
Weiß, Helmut. 2016. “Doubly-filled COMP.” *SyHD-atlas.* Accessed August 19 2023. https://www.syhd.info/apps/atlas/index.html#doubly-filled-comp.
Willkowei, Helge. 2005. ““Tante Taline sitzt schon in der Stube zu warten, dass Sie kommen.” Progressiver Aspekt im ostfriesischen Niederdeutsch.” In *Die Konstanz des Wandels im Niederdeutschen. Politische und historische Aspekte einer Sprache*, edited by Birte Arendt and Enrico Lippmann, 139–165. Hamburg: Kovač.
Wöllstein, Angelika, and Dudenredaktion , eds. 2016. Duden. Die Grammatik. Unentbehrlich für richtiges Deutsch. 9., vollständig überarbeitete und aktualisierte Auflage. Berlin: Duden.
Zwart, Cornelius Jan-Wouter. 2011. *The Syntax of Dutch.* Cambridge: Cambridge University Press.
Åkerberg, Bengt. 2012. Älvdalsk grammatik. Under medverkan av Gunnar Nyström. Mora: Ulum Dalska.